

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/208,140DATE: 12/24/1998  
TIME: 09:11:31

Input Set: I208140.RAW

This Raw Listing contains the General Information  
Section and up to first 5 pages.

*new format***ENTERED**

1 <110> APPLICANT: Del Vecchio, Alfred  
2 <120> TITLE OF INVENTION: HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN  
3 AND METHODS THEREOF TO IDENTIFY ANTIVIRAL COMPOUNDS  
4 <130> FILE REFERENCE: P50743  
5 <140> CURRENT APPLICATION NUMBER: US/09/208,140  
6 <141> CURRENT FILING DATE: 1998-12-09  
7 <160> NUMBER OF SEQ ID NOS: 34  
8 <170> SOFTWARE: FastSEQ for Windows Version 3.0  
9 <210> SEQ ID NO 1  
10 <211> LENGTH: 1710  
11 <212> TYPE: DNA  
12 <213> ORGANISM: Viral  
13 <400> SEQUENCE: 1  
14 tcaatgtctt atacctggac aggcgcactc gtcaccccg ggcgtgcgga agaacaaaaa 60  
15 ctgcccattca acgcactgag caactcgttg ctacgccatc acaatctggt atattccacc 120  
16 acttcacgca gtgcttgcca aaggcagaag aaagtcacat ttgacagact gcaagttctg 180  
17 gacagccatt accaggacgt gctcaaggag gtcaaaagcag cggcgtaaaa agtgaaggct 240  
18 aacttgctat ccgtagagga agcttgacgc ctgacgcccc cacattcagc caaatccaag 300  
19 tttggctatg gggcaaaaaga cgtccgttgc catgccagaa aggcgtagc ccacatcaac 360  
20 tccgtgtgga aagaccttct ggaagacagt gtaacaccaa tagacactat catcatggcc 420  
21 aagaacgagg tcttctgcgt tcagcctgag aaggggggtc gtaagccagc tcgtctcatc 480  
22 gtgttccccg acctgggcgt gcgcgtgtgc gagaagatgg ccctgtacga cgtggttagc 540  
23 aaactcccc tggccgtgat ggggaagctc tacggattcc aatactcacc aggacagcgg 600  
24 gttgaattcc tcgtgcaagc gtggaagtc aagaagaccc cgatggggtt cccgtatgat 660  
25 acccgctgtt ttgactccac agtcactgag agcgacatcc gtacggagga ggcaatttac 720  
26 caatgttgtg acctggaccc ccaagccgc gtggccatca agtccctcac tgagaggctt 780  
27 tatgttgggg gccctcttac caattcaagg ggggaaaact gcggctatcg cagggtgccg 840  
28 gcgagcggcg tactgacaac tagctgtggt aacaccctca cttgctacat caaggcccgg 900  
29 gcagcccgtc gagccgcagg gctccaggac tgcacatgc tcgtgtgtgg cgacgactta 960  
30 gtgcttatct gtgaaagtgc gggggtccag gaggacgcgg cgagcctgag agcctttacg 1020  
31 gaggctatga ccaggctact cgcccccccc ggggaccccc cacaaccaga atacgacttg 1080  
32 gagcttataa catcatgctc ctccaacgtg tcagtcgccc acgacggcgc tggaaaaagg 1140  
33 gtctactacc ttacctgta ccctacaacc cccctcgcga gagccgcgtg ggagacagca 1200  
34 agacacactc cagtcaattc ctggctaggc aacataatca tgtttgcccc cacactgtgg 1260  
35 gcgaggatga tactgatgac ccatttcttt agcgtcctca tagccaggga tcagcttgaa 1320  
36 caggctctta actgtgagat ctacgcagcc tgctactcca tagaaccact ggatctacct 1380  
37 ccaatcattc aaagactcca tggcctcagc gcatttttac tccacagtta ctctccaggt 1440  
38 gaagtcaata ggggtggcgc atgcctcaga aaacttgggg tcccgccctt gcgagcttgg 1500  
39 agacaccggg cccggagcgt ccgcgctagg cttctgtcca ggggaggcag ggctgccata 1560  
40 tgtggcaagt acctcttcaa ctgggcagta agaacaaagc tcaaaactcac tccaatagcg 1620  
41 gccgctggcc ggctggactt gtccggttgg ttcacggctg gctacagcgg gggagacatt 1680  
42 tatcacagcg tgtctcatgc ccggccccgc 1710  
43 <210> SEQ ID NO 2  
44 <211> LENGTH: 591

PAGE: 2

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/208,140

DATE: 12/24/1998  
TIME: 09:11:31

Input Set: I208140.RAW

```

45 <212> TYPE: PRT
46 <213> ORGANISM: Viral
47 <400> SEQUENCE: 2
48 Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala
49 1 5 10 15
50 Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg
51 20 25 30
52 His His Asn Leu Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg
53 35 40 45
54 Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp Ser His Tyr
55 50 55 60
56 Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ala Ser Lys Val Lys Ala
57 65 70 75 80
58 Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser
59 85 90 95
60 Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg Cys His Ala
61 100 105 110
62 Arg Lys Ala Val Ala His Ile Asn Ser Val Trp Lys Asp Leu Leu Glu
63 115 120 125
64 Asp Ser Val Thr Pro Ile Asp Thr Thr Ile Met Ala Lys Asn Glu Val
65 130 135 140
66 Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile
67 145 150 155 160
68 Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met Ala Leu Tyr
69 165 170 175
70 Asp Val Val Ser Lys Leu Pro Leu Ala Val Met Gly Ser Ser Tyr Gly
71 180 185 190
72 Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Gln Ala Trp
73 195 200 205
74 Lys Ser Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe
75 210 215 220
76 Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu Glu Ala Ile Tyr
77 225 230 235 240
78 Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala Ile Lys Ser Leu
79 245 250 255
80 Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser Arg Gly Glu
81 260 265 270
82 Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser
83 275 280 285
84 Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg Ala Ala Cys Arg
85 290 295 300
86 Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu
87 305 310 315 320
88 Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp Ala Ala Ser Leu
89 325 330 335
90 Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp
91 340 345 350
92 Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser
93 355 360 365
94 Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu

```

PAGE: 3

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/208,140

DATE: 12/24/1998  
TIME: 09:11:31

Input Set: I208140.RAW

95		370		375		380	
96	Thr	Arg	Asp	Pro	Thr	Thr	Pro
97	385				390		395
98	Arg	His	Thr	Pro	Val	Asn	Ser
99				405		410	
100	Pro	Thr	Leu	Trp	Ala	Arg	Met
101			420			425	
102	Leu	Ile	Ala	Arg	Asp	Gln	Leu
103			435			440	
104	Gly	Ala	Cys	Tyr	Ser	Ile	Glu
105		450				455	
106	Arg	Leu	His	Gly	Leu	Ser	Ala
107	465				470		475
108	Glu	Ile	Asn	Arg	Val	Ala	Ala
109			485			490	
110	Leu	Arg	Ala	Trp	Arg	His	Arg
111			500			505	
112	Ser	Arg	Gly	Gly	Arg	Ala	Ala
113			515			520	
114	Ala	Val	Arg	Thr	Lys	Leu	Lys
115		530				535	
116	Leu	Asp	Leu	Ser	Gly	Trp	Phe
117	545				550		555
118	Tyr	His	Ser	Val	Ser	His	Ala
119			565			570	
120	Leu	Leu	Leu	Ala	Ala	Gly	Val
121			580			585	
122							590

122 <210> SEQ ID NO 3  
123 <211> LENGTH: 1710  
124 <212> TYPE: DNA  
125 <213> ORGANISM: Viral  
126 <400> SEQUENCE: 3

127	tcaatgtctt	attcctggac	aggcgacactc	gtcaccctcg	gcgctgcgga	agaacaaaaa	60
128	ctgcccattca	acgcactgag	caactcgttg	ctacgccatc	acaatctggt	gtattccacc	120
129	acttcacgca	gtgcttgcca	aaggcagaag	aaagtcacat	ttgacagact	gcaagttctg	180
130	gacagccatt	accaggacgt	gctcaaggag	gtcaaaagcag	cggcgtcaaa	agtgaaggct	240
131	aacttgctat	ccgtagagga	agcttgacgc	ctgacgcccc	cacattcagc	caaatccaag	300
132	tttggttatg	gggcaaaaaga	cgtccgttgc	catgccagaa	aggccgtagc	ccacatcaac	360
133	tccgtgtgga	aagaccttct	ggaagacagt	gtaacaccaa	tagacactac	catcatggcc	420
134	aagaacgagg	ttttctgcgt	tcagcctgag	aaggggggtc	gtaagccagc	tcgtctcatc	480
135	gtgttccccg	acctgggctg	gcgcgtgtgc	gagaagatgg	ccctgtacga	cgtgggttagc	540
136	aagctccccc	tggccgtgat	gggaagctcc	tacggattcc	aatactcacc	aggacagcgg	600
137	gttggaattcc	tcgtgcaagc	gtggaagtcc	aagaagaccc	cgatgggggt	ctcgtatgat	660
138	accgcgtgtt	ttgactccac	agtcactgag	agcgacatcc	gtacggagga	ggcaatttac	720
139	caatgttgtg	acctggaccc	ccaagccgcg	gtggccatca	agtccctcac	tgagaggctt	780
140	tatgttgggg	gccctcttac	caattcaagg	ggggaaaact	gcggctaccg	caggtgccgc	840
141	gcgagcggcg	tactgacaac	tagctgtggt	aacaccctca	cttgctacat	caaggcccg	900
142	gcagcctgtc	gagccgcagg	gctccaggac	tgcaccatgc	tcgtgtgtgg	cgacgactta	960
143	gtcgttatct	gtgaaagtgc	gggggtccag	gaggacgcgg	cgagcctgag	agccttcacg	1020
144	gaggctatga	ccagggtactc	cgcccccccc	ggggaccccc	cacaaccaga	atacgacttg	1080

PAGE: 4

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/208,140

DATE: 12/24/1998  
TIME: 09:11:31

Input Set: I208140.RAW

```

145 gagcttataa catcatgctc ctccaacgtg tcagtcgccc acgacggcgc tggaaagagg 1140
146 gtctactacc ttaccctgta ccctacaacc cccctcgcca gagccgcgtg ggagacagca 1200
147 agacacactc cagtcaattc ctggctaggc aacataatca tgtttgcccc cacactgtgg 1260
148 gcgaggatga tactgatgac ccatttcttt agcgtcctca tagccaggga tcagcttgaa 1320
149 caggctctta actgtgagat ctacggagcc tgctactcca tagaaccact ggatctacct 1380
150 ccaatcattc aaagactcca tggcctcagc gcattttcac tccacagtta ctctccaggt 1440
151 gaaatcaata ggggtggccgc atgcctcaga aaacttgggg tcccgcctt gcgagcttgg 1500
152 agacaccggg cccggagcgt ccgcgctagg cttctgtcca gaggaggcag ggctgccata 1560
153 tgtggcaagt acctcttcaa ctgggcagta agaacaaagc tcaaactcac tccaatagcg 1620
154 gccgctggcc ggctggactt gtccggttgg ttcacggctg gctacagcgg gggagacatt 1680
155 tatcacagcg tgtctcatgc ccggcccccg 1710

```

156 <210> SEQ ID NO 4

157 <211> LENGTH: 590

158 <212> TYPE: PRT

159 <213> ORGANISM: Viral

160 <400> SEQUENCE: 4

```

161 Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr Pro Cys Ala Ala
162 1 5 10 15
163 Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser Leu Leu Arg
164 20 25 30
165 His His Asn Leu Val Tyr Ser Thr Thr Ser Arg Ser Ala Cys Gln Arg
166 35 40 45
167 Gln Lys Leu Arg His His Asn Leu Val Tyr Ser Thr Thr Ser Arg Ser
168 50 55 60
169 Ala Cys Gln Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu
170 65 70 75 80
171 Asp Ser His Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ala Ser
172 85 90 95
173 Lys Val Lys Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr
174 100 105 110
175 Pro Pro His Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val
176 115 120 125
177 Arg Cys His Ala Arg Lys Ala Val Ala His Ile Asn Ser Val Trp Lys
178 130 135 140
179 Asp Leu Leu Glu Asp Ser Val Thr Pro Ile Asp Thr Thr Ile Met Ala
180 145 150 155 160
181 Lys Asn Glu Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro
182 165 170 175
183 Ala Arg Leu Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys
184 180 185 190
185 Met Ala Leu Tyr Asp Val Val Ser Lys Leu Pro Leu Ala Val Met Gly
186 195 200 205
187 Ser Ser Tyr Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu
188 210 215 220
189 Val Gln Ala Trp Lys Ser Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp
190 225 230 235 240
191 Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu
192 245 250 255
193 Glu Ala Ile Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala
194 260 265 270

```

PAGE: 5

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/208,140

DATE: 12/24/1998  
TIME: 09:11:31

Input Set: I208140.RAW

```

195   Ile Lys Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn
196           275                      280                      285
197   Ser Arg Gly Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val
198           290                      295                      300
199   Leu Thr Thr Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg
200   305                      310                      315                      320
201   Ala Ala Cys Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys
202           325                      330                      335
203   Gly Asp Asp Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp
204           340                      345                      350
205   Ala Ala Ser Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala
206           355                      360                      365
207   Pro Pro Gly Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr
208           370                      375                      380
209   Ser Cys Ser Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg
210   385                      390                      395                      400
211   Val Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala
212           405                      410                      415
213   Trp Glu Thr Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile
214           420                      425                      430
215   Ile Met Phe Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His
216           435                      440                      445
217   Phe Phe Ser Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asn
218           450                      455                      460
219   Cys Glu Ile Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro
220   465                      470                      475                      480
221   Pro Ile Ile Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser
222           485                      490                      495
223   Tyr Ser Pro Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg Lys Leu
224           500                      505                      510
225   Gly Val Pro Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg
226           515                      520                      525
227   Ala Arg Leu Leu Ser Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr
228           530                      535                      540
229   Leu Phe Asn Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala
230   545                      550                      555                      560
231   Ala Ala Gly Arg Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser
232           565                      570                      575
233   Gly Gly Asp Ile Tyr His Ser Val Ser His Ala Arg Pro Arg
234           580                      585                      590
235   <210> SEQ ID NO 5
236   <211> LENGTH: 49
237   <212> TYPE: PRT
238   <213> ORGANISM: Viral
239   <400> SEQUENCE: 5
240   Arg Leu Leu Asp Leu Ser Ser Trp Phe Thr Val Ser Ala Gly Gly Gly
241   1                      5                      10                      15
242   Asp Ile Tyr His Ser Val Ser Arg Ala Arg Pro Arg Leu Leu Leu Leu
243           20                      25                      30
244   Gly Leu Leu Leu Leu Cys Val Gly Val Gly Ile Phe Leu Leu Pro Ala

```

PAGE: 6

**VERIFICATION SUMMARY**  
**PATENT APPLICATION US/09/208,140**

DATE: 12/24/1998  
TIME: 09:11:31

Input Set: I208140.RAW

Line ? Error/Warning

Original Text

---